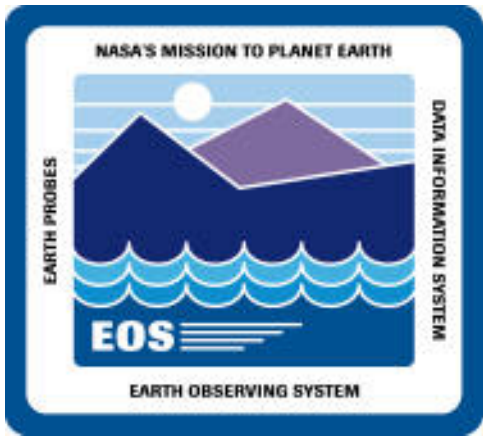


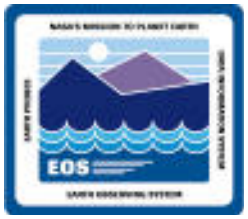
End-to-End Modeling

Nick Singer

nsinger@eos.hitc.com

19 April 1996

Agenda

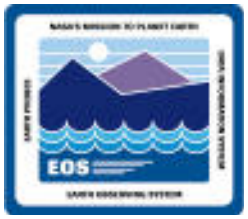


Background

Methodology Summary

Status and Output Templates

Background



Modeling activities through 1995 focused on the most critical workloads

- **Push**
- **Pull**
- **Distribution**

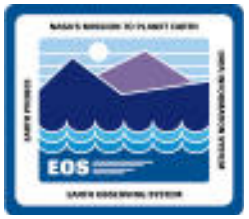
and subsystems

- **Ingest**
- **Data Server and Archive**
- **Science Processing**
- **Disks**
- **Networks**

Post-IDR-B, needed to consider all workloads & subsystems, adding

- **Infrastructure workloads (e.g. CSS/MSS activities)**
- **Science Data Server**
- **Planning**
- **Data Management**
- **Document Data Server**

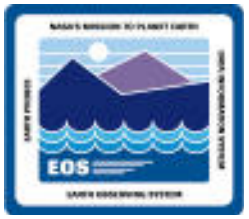
Background (cont.)



Held two Modeling Workshops

- **January 25, 1996: Modeling Assumptions and Methodology**
  - Modeling Methodology
  - Assumptions, Functions, and Parameters
  - Scenarios
  - Sources for Model Input
  - Hardware Specification
- **February 21, 1996: Interim Analyses and Results**
  - Archive Sensitivity Study
  - Sensitivity to User Pull
  - User Turnaround Times
  - Reprocessing Study
  - Failure Injection and Recovery
  - End-to-End Modeling
  - Summary and Plans for CDR and Procurement

Methodology Summary: Scope of End-to-End Modeling



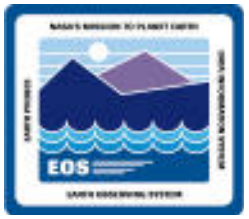
Applied to (nearly) all processing:

- **Push**
- **Pull**
- **Distribution of products**
- **Infrastructure loads**

We model processors, LANs, and disk throughput:

- **DIPHW (Distribution and Ingest Peripheral Management)**
- **ACMHW (Access Control & Management, incl. Science Data Server)**
- **SPRHW (Science Processing) *(Queuing server only)***
- **DDSHW (Document Data Server)**
- **ICLHW (Ingest Client)**
- **PLNHW (Planning)**
- **DMGHW (Data Management)**
- **DRPHW (Data Repository) *(Processors and disk only)***

Scope of End-to-End Modeling (cont.)



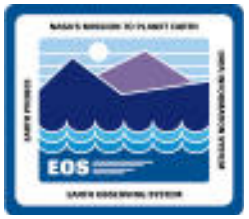
End-to-end modeling <u>doesn't</u> include:

- **V_0 loads**
- **WKSHW (Working Storage)**
- **SPRHW (Science Processing) (except for queuing server)**
- **AITHW (Algorithm Integration & Test workstations)**
- **AQAHW (Quality Assurance processing)**
- **Workstations and operator-intensive activities**
- **Archive tape hardware**
- **Disk capacity**
- **RAM**

Push (PGE) and pull processing are modeled to account for their contention for common resources

- **Calibrated to User Model, F&PRS, and dynamic BONeS model**

System Characterization



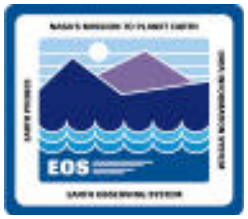
Hardware

- Machines (CPUs)
  - Sustainable MIPS
  - Number of processors per box
- Networks
  - Sustainable throughput rate (MB/sec)
  - Switch latency time
- Disks
  - Sustainable transfer rate (MB/sec)
  - Latency time

Workload

- Threads
  - Instantiation rate
  - List of activities
- Activities--use specified amounts of named resources
  - Processor
  - Network(s)
  - Disk

Threads



Acquire via Media
Acquire via Network
Activate Plan
Backup Granule
Backup List of Files
Browse Search (Motif)
Browse Search (Web)
Create 10 Day Plan
Create 3 Day Plan
Create 45 Day Plan
Delete from Pull Area
Delete from Working Storage
Estimate Cost
Execute PGE (Schedule/Queue)
Execute PGE (Stage/Process)
Generate Product On Demand
Insert L0 to Ingest
Insert Production Result to Data Server
Inventory search (Motif)
Inventory search (Web)
Monitor Pull Area
Monitor Request Queue
Monitor Working Storage
Notify following Event Trigger
Ops Intervention: Device Out of Service
Ops Intervention: Large Request
Receive Subscription Notice
Restore Backed-up Files
Retrieve (Browse)
Subset (Spatial)
Subset (Temporal)

Example Thread

Illustrative Data Only!

THREAD	EXECU-TABLE	ACTIVITY	Hardware	Network I/O (from)	Network I/O (to)	# Network Transfers	Network I/O Msg Length (MB)	Disk I/O (MB)	CPU (LOC)	# RPCs	# OF DIST. OBJ INSTS.	DBMS CALLS	CPU (MI)
Subset (Temporal)--0.0019906 per sec			Assumes temporal subsetting of single swath-based granule, nominal size of 58 MB, 50% reduction.										
	SDSRV	Request Staging Disk	ACM-3	DM/DS/CSS/MSS		2	0.00020		1,200	2	1	1	0.518
	Staging Disk Resource Manager	Allocate Disk	DRP-1						600	1			0.009
	SDSRV	Retrieve Request	ACM-3	DM/DS/CSS/MSS		2	0.00020		400	2	1	1	0.506
	Archive Resource Manger	Retrieve from archive (FSMS)	DRP-1	DM/DS/CSS/MSS		2	0.00016		400	2	1	1	0.506
	SDSRV	Perform Temporal Subsetting (HDF-EOS)	ACM-3					29.0000	10,000				0.150
	SDSRV	Update Metadata	ACM-3					0.0002	2,000			1	0.530
	Illustra	Update Metadata DB	DRP-3					0.0020				2	1.000
	SDSRV	Send Completion Status	ACM-3	DM/DS/CSS/MSS		1	0.00015		600	1	1	1	0.509

Methodology

Read characterization files; set up model

Step through threads & activities; collect statistics on load by thread and in total—by specific

- **CPU**
- **Network**
- **Disk**

Add in known background loads for each resource

Calculate utilization & expected waiting time at each resource

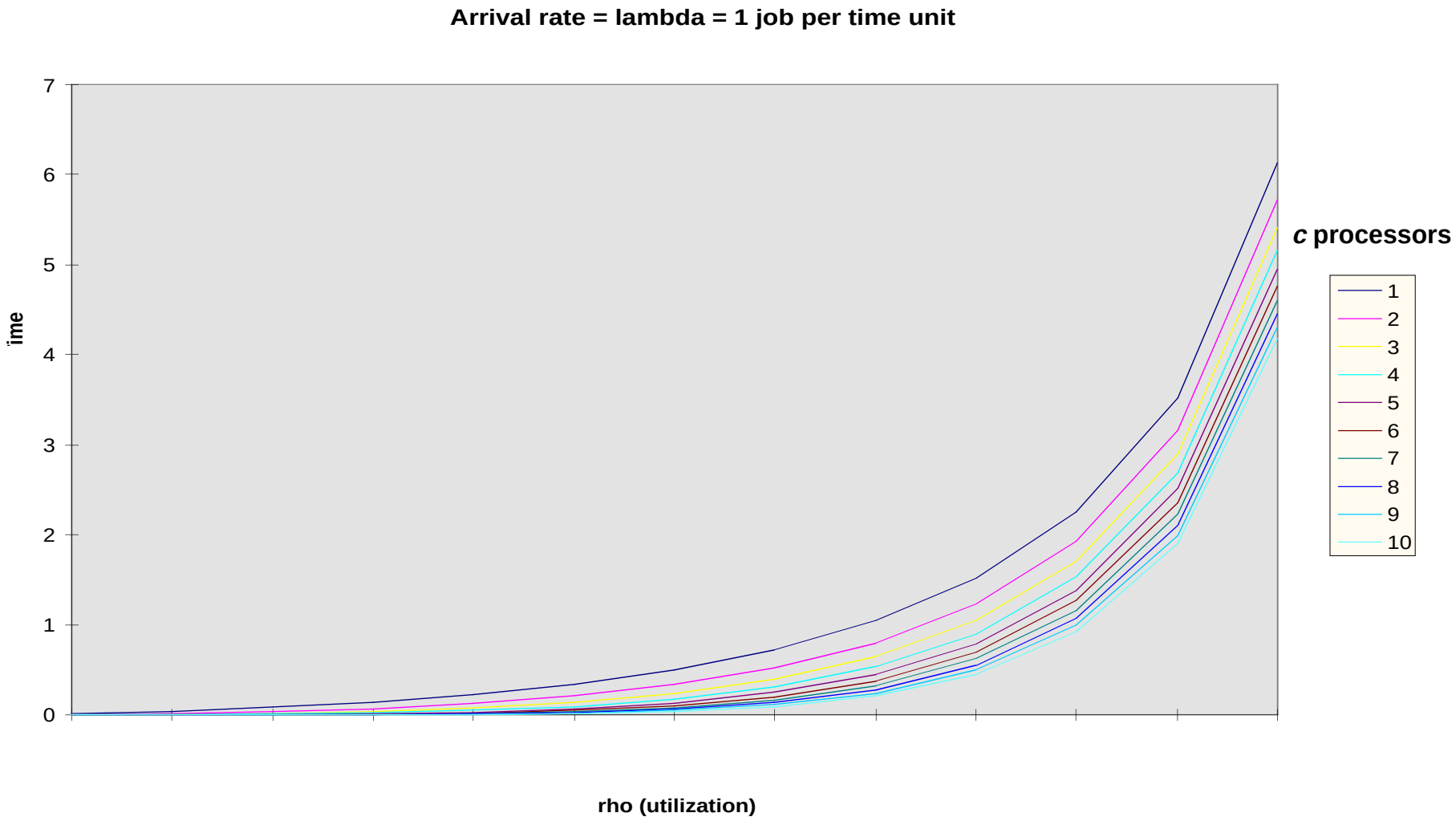
- **Calculate mean and variance of service time at each resource**
- **Utilization = ρ = Arrival rate x Service time/Number of Processors**
- **Calculate expected waiting time**

$$W_q = \frac{\lambda\, E\left[\text{service time}^2\right]}{2(1-\rho)} \ \text{(M / G / 1)} \quad \text{or} \quad W_q = \frac{(c\rho)^c \rho}{c!\lambda(1-\rho)} \Bigg/ \left(\sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!(1-\rho)} \right) \ \text{(M / M / } c\text{)}$$

Calculate end-to-end times for each thread

- **Time for a given activity at a given resource = average waiting time for the resource + service time for the activity at the resource + latency time**

Avg. Waiting Time in M/M/c Queue



Status

About 90% of threads and activities have been identified and specified

Remaining specification and data entry will take about two more weeks, followed by about three weeks of validation and calibration

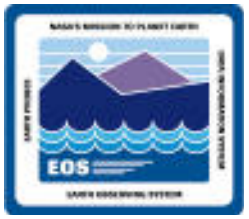
Results so far show no surprises: low loads where expected

Slides that follow illustrate the model's output templates

Template—Processors

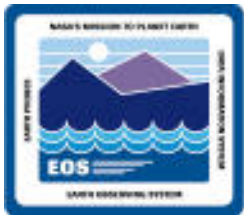
LaRC, Epoch k									
	MACHINE	ACMHW-3	DRPHW-1	DRPHW-3	DMGHW	DIPHW-1	ICLHW-1	PLNHW-1	SPRHW-7
		SDSRV	Archive FSMS	Archive DBMS (Illustra)	Data Mgmt	Distri-bution	Ingest Server	Planning Server	Queuing Server
c	**No. of Processors**	10	4	2	1	4	1	2	2
	MIPS (ea. processor)	187	187	187	146	140	187	140	140
λ	**Activity arrivals (per sec)**								
	Avg MI/activity								
	Total MIPS demand								
$1/\mu$	**Avg service time (sec)**								
	Avg [(service time)^2]								
r	**Arr rate*Avg Svce Time**								
ρ	**Avg Utilization= r/c**								
W_q	**Avg Waiting Time (sec)**								

Template—Networks



LaRC, Epoch k					
	NETWORK	PDPS	User	DM/DS/ CSS/MSS	Ingest
	MB/sec	10	10	10	10
	Switch Latency (sec)	0.010	0.010	0.010	0.010
λ	**Activity arrivals (per sec)**				
	Avg MB/activity				
	MB/sec total demand				
$1/\mu$	**Avg service time (sec)**				
	Avg [(service time)^2]				
ρ	**Arr rate*Avg Service Time**				
W_q	**Avg Waiting Time (sec)**				

Template—Disks

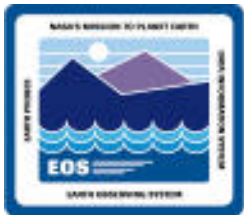


LaRC, Epoch k								
	DISK	ACMHW	DIPHW	DMGHW	DRP-DBMS	DRP-FSMS	ICLHW	PLNHW
	MB/sec	16.00	16.00	16.00	16.00	16.00	16.00	16.00
	Latency (sec)	0.00833	0.00833	0.00833	0.00833	0.00833	0.00833	0.00833
λ	**Activity arrivals (per sec)**							
	Avg MB per activity							
	Avg MB/sec							
$1/\mu$	**Avg service time**							
	Avg [(service time)^2]							
ρ	**Average utilization**							
W_q	**Average waiting time**							

Template—Thread Timings

Thread:		Subset (Temporal)			
Instantiation Rate (per sec):					
Resource	No. Visits	Avg. Waiting Time per Visit (sec)	Total Waiting Time	Total Latency + Processing	Total Time at Resource (sec)
ACM Processor	5				
DRP-FSMS Processor	2				
DRP-DBMS Processor	1				
DM/DS/CSS/MSS Net	7				
ACM Disk	3				
DRP-DBMS Disk	1				
		TOTALS			

Conclusion



End-to-End Model is a useful tool for

- **Evaluating thread timing**
- **Evaluating resource loading**
- **Performing quick what-if excursions**

End-to-End Model is easily calibrated to

- **User Model**
- **F&PRS**
- **Dynamic discrete-event simulation model**
- **Measured benchmarks**

Results will become available around early June